MBMG
Metabarcoding & Metagenomics

Metabarcoding and Metagenomics 3: 127–135
DOI 10.3897/mbmg.3.39041

Forum Paper

Adapting the (fast-moving) world of molecular ecology to the (slow-moving) world of environmental regulation: lessons from the UK diatom metabarcoding exercise

Martyn Kelly[1,2]

1 *Bowburn Consultancy, 11 Monteigne Drive, Bowburn, Durham DH6 5QB, UK*
2 *School of Geography, University of Nottingham, Nottingham NG7 2RD, UK*

Corresponding author: *Martyn Kelly* (mgkelly@bowburn-consultancy.co.uk)

Academic editor: *Agnès Bouchez* | **Received** 13 August 2019 | **Accepted** 3 December 2019 | **Published** 18 December 2019

Abstract

Development of effective metabarcoding-based tools for ecological assessment requires more than just detailed knowledge of ecology and molecular genetics. There is also a need to understand the context within which they will be used, and for the organisation that uses it to understand the techniques involved and, more especially, how the data that are produced differs from that generated by traditional ecological methods. Lessons learnt during the development of a metabarcoding tool for phytobenthos in the UK are set out in this paper. This attempted to develop a molecular "mirror" of the existing light microscopy-based approach to ecological assessment. Although this conservative approach does not exploit the full potential of metabarcoding data, it does mean that benchmarks exist against which performance and data can be judged. However, the pace of developments within molecular ecology means that regulators will need to find ways of incorporating new scientific insights whilst, at the same time, ensuring a stable regulatory regime. Installation of a metabarcoding technique within a regulatory organisation, in other words, is more than a transaction in which one approach is switched for another. A deeper transformation of the organisation is required.

Key Words

diatoms, implementation, metabarcoding, phytobenthos, Water Framework Directive

Introduction

In 2017, the Environment Agency in England took a decision to adopt a metabarcoding approach for ecological status assessment using diatoms, in order to fulfil their obligations under the European Union (EU)'s Water Framework Directive (WFD; European Union 2000). This was the first time that an approach based on Next Generation Sequencing (NGS) had been adopted by a state regulator for WFD assessments in Europe, and the decision provoked considerable interest amongst both academic ecologists and other environmental regulators. The science that underpinned this decision has been published in reports (Kelly et al. 2018a, b; 2019) but a number of other lessons were learnt during the development process that relate less to the molecular, statistical and ecological principles lying behind the new methods, and more to the "human ecology" of the organisations within which the method was to be used.

Broadly speaking, there was an unspoken assumption, made by both academic scientists involved and managers within the organisation, that implementing metabarcoding methods would involve a like-for-like replacement. A method of generating ecological classifications based on traditional morphology-based identification using light microscopy (LM) would, in other words, be replaced by one which used genomic technology (NGS). The expectation was that, with a few deft shuffles of those parts of organisational flow charts relating to how ecological data were obtained, a Brave New World of ecological assessment would be ripe for the taking.

Copyright *Martyn Kelly*. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

PENSOFT

What we know now is that the changes need to be deeper and that these may extend to the culture of the organisation itself. Key decisions, in this case, were made by experienced managers who, mostly, had little understanding of molecular biology. A lesson learnt from the UK experience is that, rather than think in terms of simple **transactions** to produce ecological data more efficiently, the organisation needs to think more boldly in terms of **transformation** and this is the subject of this essay. In particular, the current mode of implementation, with analytical work focussed in a few "High Throughput Sequencing (HTS) laboratories" raises questions about the "product" of ecological assessment. Should we think solely in terms of a metric value or status class assignment produced from a "black box"? Is there still a role for a specialist ecologist to explain and interpret results? If so, how does the organisation support and develop them? In blunt terms, the shift from traditional means of gathering ecological data to a reliance on HTS laboratories is part of a broader trend in which ecological data and evidence are gathered and processed by automated means, with the risk that professional ecologists become marginalised.

Ecological assessment requires summary information about spatially and temporally variable ecological communities in a form that can be communicated within organisations and with stakeholders in a manner that supports decision making. Most of those who use the data will not be specialists in the organisms that are being assessed; many may not even be ecologists. At its most reductionist, data are collapsed to summary metrics ("Ecological Quality Ratios", or EQRs, in the case of the WFD) which are further classified into one of five status classes. In theory, such concentrated nuggets of transferable ecological information can be reconstituted into a "guiding image" of the community or ecosystem in question (Willby 2011; Poikane et al. 2018), but only if the end-users have the necessary experience to "unlock" the information conveyed by the classification. The extent (and, indeed, the need) to which this happens in practice will vary within and between organisations, but we can assume that higher levels of management will be more focused on a face-value interpretation of classifications, and their implications for catchment management, and less engaged with the intricacies of particular organism groups.

Couple this with a trend towards metrics that are based upon complicated algorithms (Clarke et al. 2003), with supporting routines that deal with uncertainty assessment (Clarke 2013), along with computer-based decision-support tools for catchment risk assessments (e.g. Holmes et al. 2005; Berlekamp et al. 2007). It then becomes easy to make an attractive economic case for replacing time-consuming analyses of ecological samples by streamlined HTS laboratories linked directly to sophisticated data analysis programs. If, within this dystopian scenario there is little space for an ecologist's experience to influence the decision-making process then it is a short step to assessments based on taxonomy-free approaches (advocated by Pawlowski et al. 2018, amongst others), at which point the link with the wider knowledge base is lost. It is, I have realized, easy for middle and senior management to be seduced by the promise of these new technologies, without realizing the scale of what might be lost.

Engagement with one of the first instances in which an ecological metabarcoding approach was taken right through from conception to operational deployment offers some perspectives on the potential of this method. Whilst there is no shortage of academic papers hyping the potential of genomic approaches (Baird and Hajibabaei 2012; Woodward et al. 2013; Leese et al. 2018) nor of position papers written by managers inside government agencies, there are, inevitably, issues that arise during implementation that could not have been foreseen. Philosophers recognise that, whilst the type of rational argument that these papers encapsulate is necessary, it needs to be moderated and qualified by practical experience (so-called 'standpoint epistemology': Lukács 1971) and it is in this spirit that this essay was written. Whilst primarily focussed on metabarcoding in the context of European environmental legislation, the broader lessons should be transferable to other parts of the world, and to other emerging ecogenomic technologies, too.

Political context

Whilst the WFD sets out a broad ambition for sustainable use of aquatic resources within the EU, it is the responsibility of each individual Member State to implement the Directive within their territory as they see fit. The result is a wide diversity of methods (Birk et al. 2012; Kelly et al. 2014; Poikane et al. 2017) often rooted in historical traditions within a country (Kelly et al. 2016). There are also differences in the structure of environmental regulators. Some of the decisions made in the UK, as a result, may make less sense in countries where environmental regulation is organised differently.

The implementation of the metabarcoding system in England also needs to be viewed within a broader political context. In particular, responsibility for the environment is devolved to the four constituent members of the United Kingdom: England, Scotland, Wales and Northern Ireland. Political and policy control rests with elected representatives but day-to-day regulation is the responsibility of agencies that are answerable to the administrations, but which retain some flexibility in the way that they manage the environment. At the same time, when dealing with the European Union on issues relating to the WFD, the United Kingdom presents a single face, via a semi-independent body, the UK Technical Advisory Group for the Water Framework Directive (UK TAG: http://www.wfduk.org). Whilst UK TAG often works smoothly, there are cases (this was one) when tensions amongst the national environment agencies can arise.

Finally, the development of the UK metabarcoding approach needs to be understood against the background of the financial crisis that engulfed the UK (and much

of the rest of the world) in 2009. This had major implications for the financing of the public sector in the UK, with environmental regulation being one area that had to operate with reduced budgets. This, in turn, led to a desire to seek economies whilst still fulfilling their legal obligations with respect to EU legislation. The internal buzzword in the Environment Agency was "more for less"; in practice, several areas of activity (including monitoring) were reduced and searches for more efficient means of collecting environmental data were initiated. Claims that metabarcoding would reduce costs compared to conventional ecological analyses (Baird and Hajibabaei 2012) raised the possibility that this would allow effective monitoring with a significantly lower budget. A preliminary review of the potential of molecular techniques for ecological monitoring (Kille et al. 2011) identified diatoms as a group that could be used to establish the viability of metabarcoding approaches.

By comparison, despite dominating political debate in the UK for the past three years and generating considerable uncertainty, Brexit had a relatively small effect on the project. Following the referendum, the UK environment agencies adopted a "business as usual" approach with regard to their responsibilities to the European Commission, with the expectation that there would be a transition period between formally leaving the EU and full political independence. In any case, the principles of the WFD have been transposed to UK law and will continue to be applied for the foreseeable future, albeit outside the jurisdiction of bodies such as the European Court of Justice.

The WFD has had two profound effects on the way ecological assessment was performed in UK environment agencies: one positive and one negative. The positive effect is that the requirement to evaluate a broad range of ecosystem components initially encouraged the diversification of biological skills amongst ecologists. In 2010, for example, there were about 40 biologists distributed throughout the UK environment agencies with diatom identification skills, most of whom were also able to analyse invertebrate samples and perform macrophyte surveys as well as responding to a wide range of enquiries about local phenomena from the public. They collected their samples from a limited geographical area that they grew to know well and, over time, developed a holistic awareness of freshwater systems in this area. On the other hand, the sophisticated nature of the calculations which underpin WFD assessments, along with the need for a harmonised approach to regulation within the country, has contributed to the centralisation of decision-making within the agencies and an increasing reliance on standardised assessment tools in place of local ecological expertise.

Overview of UK diatom metabarcoding project

Following the preliminary literature review (Kille 2011) and a brief feasibility study, the first phase of the work (based around the Roche 454 sequencing platform) resulted in the development of a barcode library but primers subsequently had to be re-designed to fit the requirements of the Illumina Mi-Seq sequencing technology after withdrawal of support for the Roche 454. Subsequent work (summarised in Table 1) focused on development of a version of the TDI that was optimised for metabarcoding data and understanding differences between LM and NGS outputs.

Once these basic components of a metabarcoding system were in place attention shifted to implementation and implications for regulation. The means for computing "expected" values of the TDI (i.e. the denominator in Ecological Quality Ratio calculations) was recognised as a weakness of the LM approach used in UK and the metabarcoding project presented an opportunity to re-examine this, resulting in an improved reference model based on stronger conceptual foundations (Kelly et al. 2018). This also precipitated a new look at how phytobenthos and macrophyte data should be combined to provide an integrated assessment of "macrophytes and phytobenthos", as required by the WFD. However, by the time this work was completed, planning for the third cycle of the WFD was too far advanced for a wholesale shift in the approaches to assessment to be contemplated and this was not adopted.

Table 1. Outputs from the UK diatom metabarcoding project.

Output	Reference and notes
RbcL primer optimised for Illumina Mi-Seq	Kelly et al. (2018a)
Bioinformatic pipeline (QIIME2)	Kelly et al. (2018a)
Barcode library (1232 strains / 346 taxa)	Kelly et al. (2018a); Barcodes are now freely-accessible in the Diat.barcode database (Rimet et al. 2019)
Optimised version of TDI for LM (TDI5LM) and recalibrated version of TDI for metabarcoding (TDI5NGS)	Principle and method described in Kelly et al. (2018a); final iteration in Kelly et al. (2018c).
Comparison of reproducibility and repeatability for LM and NGS at different scales	Kelly et al. (2018a)
Guidance on sampling strategies to minimise contamination when sampling diatoms for metabarcoding	Kelly et al. (2018b)
Preliminary guidance on interpreting results from diatom metabarcoding analyses	Annex to Kelly et al. (2018b)
Revised reference models for estimating ecological status using diatoms (applicable to both LM and NGS versions of TDI)	Kelly et al. (2018c)
Revised combination rules for combining macrophytes and phytobenthos for ecological status assessments	Kelly et al. (2018c)

What did we learn?

Lesson 1: be clear about the "product"

Ecological assessment produces three separate "products":

1. The nuggets of transferable information (EQRs and ecological status classifications in the case of the WFD);
2. Underpinning data (lists of taxa and their relative abundances); and,
3. Collateral benefits, as engagement with samples and surveys increases an individual's knowledge and understanding of the ecosystems under consideration.

The focus of the UK phytobenthos metabarcoding project was primarily on the first of these. Most of those managing the projects had "head office" functions with responsibility for national classifications and strategy, and on reporting UK compliance. Although taxa lists played an important role for interpreting assessment results in the past, their primary role for statutory assessment is as raw material for metric calculation. Budget constraints have led to a greater focus on standardized approaches to reporting and risk assessment with fewer opportunities for local staff to interpret diatom outputs.

There are substantial differences in the relative abundance of diatom species in lists produced by conventional and metabarcoding approaches (Vasselon et al. 2017; Kelly et al. 2018). The UK response was to produce a recalibrated metric optimized for metabarcoding data: an acceptable 'fix' if the focus of assessment is the final metric value but with implications for data interpretation (see below).

The ability of metabarcoding to detect species is dependent, to a large extent, on the quality of barcode libraries (Weigand et al. 2019). For diatoms, these currently include only about 20% of known freshwater species, with a bias towards common taxa (Rimet et al. 2019). Extending taxa dictionaries to encompass all microalgae would be prohibitively expensive so there is a need for an informed conversation about alternative approaches. The extreme position is 'taxonomy-free' approaches (Apothélez-Perret-Gentil et al. 2017), which offer no opportunity to linking to a wider knowledge base but there are many intermediate options. When considering these, however, the WFD's emphasis on stakeholder engagement (Article 14: "Public information and consultation") also needs to be considered. The taxa list should provide a bridge between the evidence upon which classifications are based and an understanding of the ecosystem which, in turn, can be communicated beyond a narrow group of specialists.

Both the absence of taxa from barcode libraries and difference in their representation in light microscopy and metabarcoding data need to be understood by those engaged in interpretation. This is important because, although classifications are broadly comparable, it is possible that individual sites within a data cloud may be classified differently using light microscopy and metabarcoding (see below). Responsibility for explaining such shifts often falls to local staff who will, in turn, need to understand and explain the reasons, for which a list of taxa and abundance assessments are necessary. The taxa list also provides a link to a wider pool of knowledge, including ecological traits (Rimet and Bouchez 2012; Tapolczai et al. 2016) which would enable a local ecologist to interpret changes in a community over time. Thus, the implementation of a metabarcoding program needs to be accompanied by a structured training program accessible to all who might need to interpret the data, to ensure that they are aware of how metabarcoding differs from the traditional data with which they are familiar. However, to date, there has been relatively little focus from management on helping local ecologists understand differences between the traditional and molecular data.

Whilst the previous paragraph ended with a call for a "structured training program" as part of the transition to metabarcoding, this one starts by lamenting the decline of a widescale "unstructured training programme" arising from the engagement of individual ecologists in the UK with all stages of assessment, from sampling to interpretation, within a relatively small geographic area often for several years. A business strategy that prioritises 'efficiency' over depth of technical knowledge and resilience in its experts has led to fragmentation of the data chain in recent years. This, in turn, leaves ecologists increasingly frustrated in their desire to develop the detailed local ecological understanding needed to properly assess and interpret environmental data. Because sampling phytobenthos for metabarcoding is a straightforward process that can be delegated to non-specialist technicians, after which all aspects of metabarcoding analysis take place in remote laboratories, an ecologist's first encounter with a sample may be as a list of taxa in a spreadsheet, or even as a final metric or classification result. As a result, it is increasingly difficult for ecologists to contextualise data as they are no longer so familiar with the locality from which the data are derived. The implications of this may take a few years to become apparent as the first generation of ecologists who interpret metabarcoding data will have this experience on which to draw. After that, will there be any capacity for an informed "reality check" of the outcomes of bioinformatic pipelines?

Lesson 2: learn to walk before you try to run

Hering et al. (2018) reviewed implementation options for metabarcoding in Europe and identified two broad strategies:

- Option 1: Like-for-like replacement of existing methods. In other words, use NGS as alternative means of data acquisition, but continuing with existing principles behind metrics, reference conditions and status class boundaries;

- Option 2: Adopt the "Biomonitoring 2.0" philosophy propounded by Baird and Hajibabaei (2012), developing a new generation of indices using the full potential of NGS data to explore the greater taxonomic depth that is available, a wider range of organisms and, potentially, to develop metrics that measure ecosystem function as well as structure.

The UK phytobenthos method fits "option 1". When development started there were no studies that had used NGS to evaluate the response of diatom assemblage composition along environmental gradients so the foundation of evidence that existed for LM studies provided a strong hypothesis for determining what taxa we should expect to detect in our metabarcoding output. In brief, if a species is present in a LM analysis, you should expect to find it in NGS output. If not, then you should take the opportunity to understand why this is not the case. This hypothesis also prompted us to ask questions about quantification (why is a particular taxon less abundant in NGS output than in an LM analysis, and vice versa?) and the breadth of our taxa dictionary. The opposite also applies: if a species is present in NGS output, then does this mean that it has previously been overlooked using traditional analyses (Blackman et al. 2017)? Knowing that some taxa abundant in metabarcoding data were absent from light microscopical analyses of the same sample also allowed us to track down some faults in the laboratory and bioinformatic stages. Having learnt how NGS output relates to the existing knowledge base, we are in a stronger position to move forward and start to exploit the potential of NGS more fully.

A further point is that a project that sets out to develop a practical tool for ecological assessment is constrained by the wording of current legislation and regulations that it will be used to enforce. In the case of the WFD, the normative definitions for ecological status refer to "composition" and "abundance" and Member States are required to intercalibrate their methods in order to ensure that all share a common level of ambition with respect to key WFD objectives (Birk et al. 2013). This, too, should be more straightforward if the first generation of metabarcoding methods produce output that can be readily related to that from neighbouring countries.

On the other hand, the UK experience has, to date, not yielded a method that is appreciably better than the existing one, leading some to ask whether our conservative "option 1" approach was justified. In particular, uncertainties associated with both LM and NGS approaches mean that the classification of some water bodies will change solely as a result of the switch in method (see below). Taking a longer perspective, the current NGS method has laid foundations upon which methods that better exploit the potential of NGS can be developed. However, at the time of writing, no-one has proposed a strong assessment concept for phytobenthos that could improve upon, and potentially replace, the present "option 1" implementation.

Lesson 3: acknowledge (and address) your ignorance

Moving from methods based on identifying organisms based on their morphology to identification based on genetic traits involves a major paradigm shift, not just in our taxonomic understanding, but also in how data are obtained, and their ecological meaning. Most of those responsible for making decisions about the implementation of metabarcoding trained in the era of morphology-based taxonomy and ecology, and had little or no understanding of the laboratory and bioinformatic stages involved in producing NGS output. Conversely, those responsible for primer design, DNA extraction and PCR, high-throughput sequencing and development of the bioinformatics pipeline were excellent molecular biologists but had little awareness of phytobenthos ecology or the regulatory framework within which the method would be used.

A further complication was that diatom taxonomy is also in the midst of its own paradigm shift, leading to reviews of species and generic limits (Ruck et al. 2016), the description of many new taxa (e.g. Kahlert et al. 2019) and a better understanding of just how much we don't know (Mann and Vanormelingen 2013). Molecular genetics has played a major role in this process and, indeed, informed the initial choice of a barcode for the metabarcoding tool (Mann et al. 2010). However, once again, appreciating the scale and implications of this was a challenge for managers (mostly from disciplines where taxonomy was more settled) responsible for determining the course of metabarcoding implementation.

Development of metabarcoding-based methods for ecological assessment inevitably calls for multidisciplinary teams, with individuals playing to their strengths. There is also a need for a measure of mutual understanding: awareness of how NGS data are produced by regulators and "old school" ecologists, awareness of the underlying legal context by bioinformaticians, awareness of the "natural history" of organisms by those whose focus is either high-level catchment management and nationwide strategy or, on the academic side, analysis of "big data". With the benefit of hindsight there was not enough overlap between the skills of the different specialists involved in the development of the UK diatom metabarcoding tool, and some basic education in each other's perspectives would have been fruitful. This was particularly obvious at later stages when a wider pool of individuals needed to be engaged yet where, due to pressures of time and budget, there were more teleconferences and fewer face-to-face meetings. The former, typically constrained to no more than a couple of hours, gave scant opportunity to understand and learn.

Lesson 4: finding the balance between stability and innovation

There should be a general expectation that the process of environmental regulation is stable and, when a change does happen, we are given due warning and explanation.

A utility company, for example, may spend some years evaluating the case for, and design of an improvement to a wastewater treatment plant, with an expectation that the investment would be recouped over a set time period. They would not anticipate, for example, the regulator adjusting targets during that period due to a change in the way that ecological status was evaluated. When a change does occur, they would expect this to be supported by sound evidence.

Yet genetic technologies promise greatly enhanced capacities to evaluate the environment and, in these early years of implementation, we are learning all the time. Is it not in the public interest to incorporate this new knowledge into regulatory regimes as soon as possible? Finding the balance between the competing demands for "stability" and "improvement" (and, in the latter case, sifting the genuinely useful from the over-hyped) is going to be one of the major challenges facing those charged with managing the transition to molecular approaches for ecological assessment over the next decade and beyond.

Two extreme views expressed during this project were that molecular methods should not be adopted until their development had "stabilised" or that they should be adopted as soon as they were fit-for-purpose but be "locked down" so that regulation was based on a single, unambiguous implementation of the method. The former argument is not tenable because of the rapid evolution of metabarcoding technologies. When development of the UK diatom metabarcoding method started, the Roche 454 was the state-of-the-art platform for NGS and primers were designed around its capabilities. During the course of the project, this machine became obsolete and primers had to be re-designed to meet the requirements of the Illumina Mi-Seq. There is no reason to assume that this, too, might follow the Roche 454 into obscurity at some point and that the method will need to be fitted to a new platform. There are, in addition, exciting new technologies available, such as Oxford Nanopore's Minion (Grey et al. 2015), which might well make molecular analysis more accessible to field-based ecologists without the need to engage a remote HTS laboratory. The argument for waiting for methods to stabilise is simply not realistic in the era of molecular ecology.

In practice, the relationship between phytobenthos assessments based on LM and NGS is not perfect (see Kelly et al. 2018b) and some sites classified as 'moderate status' using LM are classified as 'good status' using NGS, or vice versa. Systematic elements of this variation can be minimised to some extent (e.g. by calibrating the NGS method to fit established boundaries) and it may also be possible to make a case that the new method is a genuine improvement over the old (e.g. if the relationship with pressure is stronger, or the underlying model for predicting 'expected' metric values is improved). However, there is also an element of random variation, with sites shifting position within the data cloud for no explainable reason. This is, remember, no longer an academic exercise of hypothesis testing in the face of noisy data, but a case where a regulator and a utility company may have developed future investment plans using a scenario that they are now being told is wrong.

At the heart of this, lies the WFD's controversial 'one out, all out' rule by which the final status of a water body is determined by the lowest status of any of the measured components of ecological status (Carvalho et al. 2018). In practice, a change in overall status due to a switch to NGS should only occur in those situations where phytobenthos either was, or becomes, the trigger for a downgrade. In many cases, assessments of other organism groups will either confirm or overrule any change that is purely due to the shift in phytobenthos classification. But this will still leave some situations where a change in classification may occur due to the method switch. Sites will be more vulnerable to this if the number of criteria monitored has been pared back to save money, but also where the number of samples contributing to a phytobenthos assessment is small (leading to high standard error and, thereby, uncertainty).

After this initial jolt to the classification regime, we should expect a period of incremental improvements in the science and the next question is whether these should be incorporated into assessments or not. The approach adopted in the UK was to take the version of the metabarcoding tool that emerged after three iterations of the prototype and to "lock down" all aspects of procedure (including bioinformatics pipeline, taxa dictionary and species sensitivities) to ensure consistency in output. For some organism groups this may not be a problem but, given the rate at which diatom taxonomy is developing (see lesson 3), adherence to a fixed list soon diverges from "best practice". "Locking down" the taxa dictionary and bioinformatics may seem like a means of ensuring stability but, given that over 40% of reads cannot be assigned to sequences in the barcode library, this also acknowledges a large level of uncertainty in outputs. At the point at which the UK barcode library was "locked down", the rate of unassigned reads was slightly lower than that obtained by French colleagues; two years later, they (who have continued to develop their barcode library and bioinformatics) have halved their rate of unassigned reads (author, Rachel Glover, Frédéric Rimet, Valentin Vasselon, unpublished data).

The decision to "lock down" the method in the name of regulatory stability could well prove to be the biggest mistake in the implementation process. A general principle in quality management is that variation that can be controlled should be controlled whilst that which cannot be controlled should be quantified and incorporated into the decision-making process. In effect, "fixing" the metabarcoding procedure treats a source of potentially controllable variation as uncontrollable. Filling a few gaps in the barcode library is unlikely to have a noticeable effect on the overall strength of fit between the LM and NGS methods, or between the NGS method and principal pressure gradients. However, it might correct the position of a few points within the data cloud and result in better regulation at the locations these represent.

Would this be at the expense of the "stability" which is, clearly, an important requirement of a regulatory regime? In theory (and depending on the means by which key regulatory boundaries are set), any alteration to the means by which a metric is produced could have implications for regulation. In practice, however, such alterations are likely to be minor and, in any case, methods already exist to compensate for this. Over the course of WFD implementation, the intercalibration process has ensured that Member States all share a similar level of ambition when setting ecological status boundaries. These have generated a set of methods (Birk et al. 2013; Poikane et al. 2015) designed to ensure spatial compatibility (i.e. between Member States) but which could also be used to ensure temporal compatibility (within a Member State over time). Indeed, Birk et al. (2015) was written with this express purpose in mind. These should allow the position of status class boundaries to be held constant whilst the underlying metrics are improved. This will, of course, need to be managed with great care and with good version control and high transparency to ensure regulatory stability.

Lesson 5: Work with European colleagues

The credibility of methods is strengthened when there is broad international agreement, underpinned by international standards – endorsed either by the International Standards Organisation (ISO) or Comité European de Normalisation (CEN). Conversely, the absence of such a standard may be seen as a barrier to implementation. In the case of diatoms, informal discussions with colleagues performing similar work elsewhere in Europe led to an approach to the relevant technical committee of CEN and, eventually, publication of two Technical Reports (CEN 2018a, b). These are not full-fledged "standards" because there are still important knowledge gaps (e.g. in preservation of DNA over long time periods) but are steps on a longer journey towards full harmonisation of methods. The COST project, DNAqua-net (Leese et al. 2018) has, in turn, provided logistical support for performing some method comparisons which will, over time, contribute to this harmonisation.

A second area where international co-operation has been invaluable has been in the development and curation of a barcode library (Rimet et al. 2019). The importance of barcode libraries has already been stressed (see Weigand et al. 2019): the scale of the task, particularly for species-rich phyla, is such that no single country has enough specialists or resources to complete the task. In view of the limited resources available, there seems to be little sense in each country developing a separate barcode library.

Conclusion: transaction or transformation?

At face value, the outcome of five years research was a metabarcoding assessment tool that had a similar level of performance to the light microscopy-based tool that it replaced, at a slightly lower cost per sample. For internal reasons, some components of the new system (e.g. a revised means for computing "expected" values of metrics: Kelly et al. 2018b) were not adopted by the regulators, so there was little obvious benefit to offset the disruption to existing classifications (see lessons 2 and 4). Consequently, whilst environment agencies in England and Scotland seemed happy to adopt the method, the UK political administrations eventually decided not to adopt the NGS method. Given the intra- and inter-agency rancour that accompanied the final stages of the development of this method, has it all been worthwhile?

Either the glass is half full or it is half empty: the optimistic interpretation is that, although the phytobenthos metabarcoding method is no better than the approach it replaced, having taken these first steps into the unknown, the UK's environment agencies are now better placed to exploit metabarcoding technologies, not just for phytobenthos but for other groups of organisms too. The lessons learnt are partly technical but also wider, addressing how large organisations make decisions about emerging technologies.

It is, however, relatively easy to be cynical about how large and ponderous government agencies respond to new technology. The best we can hope for is that the lessons above are fuel for thought so that mistakes are not repeated. Once again, this is less to do with the intricacies of molecular biology and more to do with managing change, which is never an easy task. It is easier, perhaps, to recognise the need for change in others than it is to examine one's own shortcomings, so I end with a personal "mea culpa": I entered this process with 20 years' experience of using phytobenthos to address questions raised by environmental legislation but largely ignorant of molecular ecology. "Lesson 3" is particularly heartfelt for me: I should have gone on a course to fill gaps in my understanding about molecular technologies (and, in particular, to have been able to engage more fully with bioinformatics) much sooner. I might then have been able to play a more effective part in debates about how to handle barcodes that could not be linked to diatom sequences in the barcode library.

A second "wish" that was out of my hands would have been to insist on a phased introduction of the new method, rather than a single blanket imposition across England in 2017, along with close interaction with the first cohorts of ecologists handling the end-products in order to improve interpretation skills and have a conversation about where the method was working well or less well.

And, one final "wish" would be that the UK had not decided to leave the European Union at the time these decisions about metabarcoding were being made. Collaboration with scientists from around the EU (not least via DNAqua-net) has proved to be invaluable, partly as their research agendas often helped answer questions that our administrators deemed unworthy of funding, and partly because the scrutiny of EU institutions provides a valuable counterbalance to the flaws of domestic environmental policy.

Acknowledgements

Many thanks to Rosetta Blackman (EAWAG), Richard Chadd (Environment Agency) and Tim Jones (Environment Agency) for comments on drafts and Pieter Boets (Universitiet Ghent) and Estelle Lefrançois for reviews of the manuscript. Thanks, too, to Kerry Walsh (Environment Agency) who steered the early development of the diatom metabarcoding approach in England. This paper was developed from talks given at two workshops funded by the COST Action DNAqua-Net (CA15219), supported by the COST (European Cooperation in Science and Technology) program.

References

Apothéloz-Perret-Gentil L, Cordonier A, Straub F, Iseli J, Esling P, Pawlowski J (2017) Taxonomy-free molecular diatom index for high-throughput eDNA biomonitoring. Molecular Ecology Resources 17: 1231–1242. https://doi.org/10.1111/1755-0998.12668

Baird DJ, Hajibabaei M (2012) Biomonitoring 2.0: A new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. Molecular Ecology 21: 2039–2044. https://doi.org/10.1111/j.1365-294X.2012.05519.x

Bennett C, Owen R, Birk S, Buffagni A, Erba, S, Mengin N, Murray-Bligh J, Ofenböck G, Pardo I, van de Bund W, Wagner F, Wasson JG (2011) Bringing European river quality into line: an exercise to intercalibrate macro-invertebrate classification methods. Hydrobiologia 667: 31–48. https://doi.org/10.1007/s10750-011-0635-2

Berlekamp J, Lautenbach S, Graf N, Reimer S, Matthies M (2007) Integration of MONERIS and GREAT-ER in the decision support system for the Geran Elbe river basin. Environmental Modelling and Software 22: 239–247. https://doi.org/10.1016/j.envsoft.2005.07.021

Birk S, Bonne W, Borja Á, Brucet S, Courrat A, Poikane S, Solimini A, Van De Bund W, Zampoukas N, Hering D (2012) Three hundred ways to assess Europe's surface waters: An almost complete overview of biological methods to implement the Water Framework Directive. Ecological Indicators 18: 31–41. https://doi.org/10.1016/j.ecolind.2011.10.009

Birk S, Willby NJ, Kelly MG, Bonne W, Borja A, Poikane S, van de Bund W (2013) Intercalibrating classifications of ecological status: Europe's quest for common management objectives for aquatic ecosystems. Science of the Total Environment 454–455: 490–499. https://doi.org/10.1016/j.scitotenv.2013.03.037

Birk S, Willby N, Poikane S, van der Bund W (2015) Procedure to fit new or updated classification methods to the results of a completed intercalibration exercise. CIS Guidance Document No. 30, European Commission Joint Research Centre, Ispra.

Blackman RC, Constable D, Hahn C, Sheard AM, Durkota J, Hänfling B, Lawson Handley L (2017) Detection of a new non-native freshwater species by DNA metabarcoding of environmental samples – first record of *Gammarus fossarum* in the UK. Aquatic Invasions 12: 177–189. https://doi.org/10.3391/ai.2017.12.2.06

Carvalho L, Mackay EB, Cardoso AC, Baattrup-Pedersen A, Birk S, Blackstock KL, Borics G, Borja A, Feld CK, Ferreira MT, Globevnik L, Grizzetti B, Hendry S, Hering D, Kelly MG, Langaas S, Meissner K, Panagopoulos Y, Penning E, Rouillard J, Sabater S, Schmedtje U, Spears BM, Venohr M, van de Bund W, Solheim AL (2019) Protecting and restoring Europe's waters: An analysis of the future development needs of the Water Framework Directive. Science of the Total Environment 658: 1228–1238. https://doi.org/10.1016/j.scitotenv.2018.12.255

CEN (2018a) Water quality – Technical report for the management of diatom barcodes. Technical Report TR17244. Comité Européen de Normalisation, Brussels.

CEN (2018b) Water quality – Technical report for the routine sampling of benthic diatoms from rivers and lakes adapted for metabarcoding analyses. Technical Report TR17245. Comité Européen de Normalisation, Brussels.

Clarke RT, Wright JF, Furse MT (2003) RIVPACS models for predicting the expected macroinvertebrate fauna and assessing the ecological quality of rivers. Ecological Modelling 160: 219–233. https://doi.org/10.1016/S0304-3800(02)00255-7

Clarke RT (2013) Estimating confidence of European WFD ecological status class and WISER Bioassessment Uncertainty Guidance Software (WISERBUGS). Hydrobiologia 704: 39–56. https://doi.org/10.1007/s10750-012-1245-3

Dines RA, Murray-Bligh J (2000) Quality assurance and RIVPACS. In: Wright JF, Sutcliffe DW, Furse MT (Eds) Assessing the Biological Quality of Freshwaters: RIVPACS and other techniques, 71–78.

European Union (2000) Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. Official Journal of the European Communities, Series L 327, 1–73.

Gray C, Bista I, Creer S, Demars BOL, Falciani F, Monteith DT, Sun X, Woodward G (2015) Freshwater conservation and biomonitoring of structure and function: genes to ecosystems. In: Aquatic Functional Biodiversity: an Ecological and Evolutionary Perspective. Elsevier, Amsterdam, 241–271. https://doi.org/10.1016/B978-0-12-417015-5.00010-4

Hering D, Borja A, Jones JI, Pont D, Boets P, Bouchez A, Bruce K, Drakere S, Hänfling B, Kahlert M, Leese F, Meissner K, Mergen P, Reyjol Y, Segurado P, Vogler A, Kelly MG (2018) Implementation options for DNA-based identification into ecological status assessment under the European Water Framework Directive. Water Research 138: 192–205. https://doi.org/10.1016/j.watres.2018.03.003

Holmes MGR, Young AR, Goodwin TH, Grew R (2005) A catchment-based water resource decision-support tool for the United Kingdom. Environmental Modelling and Software 20: 197–202. https://doi.org/10.1016/j.envsoft.2003.04.001

Kahlert M, Kelly MG, Mann DG, Rimet F, Sato S, Bouchez A, Keck, F (2019) Connecting the morphological and molecular species concepts to facilitate species identification within the genus Fragilaria (Bacillariophyta). Journal of Phycology 55: 948–970. https://doi.org/10.1111/jpy.12886

Kelly MG, Schneider SC, King L (2015) Customs, habits, and traditions: the role of nonscientific factors in the development of ecological assessment methods. Wiley Interdisciplinary Reviews: Water. https://doi.org/10.1002/wat2.1068

Kelly M, Boonham N, Juggins S, Kille P, Mann D, Pass D, Sapp M, Sato S, Glover R, (2018a) A DNA based diatom metabarcoding approach for classification of rivers. Science Report SC140024/R, Environment Agency, Bristol.

Kelly M, Boonham N, Juggins S, Mann D, Glover R (2018b) A DNA based diatom metabarcoding approach for classification of rivers: Further development. Science Report SC160014/SR, Environment Agency, Bristol.

Kelly, M., Juggins, S, Phillips G, Willby N (2018c) Evaluation of benthic diatom classifications in UK rivers using LM and NGS methods. Report number E18-56. Scottish Environment Protection Agency, Stirling.

Kelly M, Urbanic G, Acs E, Bennion H, Bertrin V, Burgess A, Denys L, Gottschalk S, Kahlert M, Karjalainen S-M, Kennedy B, Kosi G, Marchetto A, Morin S, Picinska-Fałtynowicz J, Poikane S, Rosebery J, Schoenfelder I, Schoenfelder J, Varbiro, G (2014) Comparing aspirations: intercalibration of ecological status concepts across European lakes for littoral diatoms. Hydrobiologia 734: 125–141. https://doi.org/10.1007/s10750-014-1874-9

Kille P (2011) A review of molecular techniques for ecological monitoring. Science Report SC090010, Environment Agency, Bristol.

Leese F, Bouchez A, Abarenkov K, Altermatt F, Borja Á, Bruce, K, Ekrem T, Ciampor Jr F, Ciamporová-Zaťovičová Z, Costa FO, Duarte S, Elbrecht V, Fontaneto D, Franc A, Geiger MF, Hering D, Kahlert M, Kalamujić-Stroil B, Kelly M, Keskin E, Liska I, Mergen P, Meissner K, Pawlowski J, Penev L, Reyjol Y, Rotter A, Steinke D, van der Wal B, Vitecek S, Zimmermann J, Weigand AM (2018) Why We Need Sustainable Networks Bridging Countries, Disciplines, Cultures and Generations for Aquatic Biomonitoring 2.0: A Perspective Derived From the DNAqua-Net COST Action. Advances in Ecological Research 58: 63–99. https://doi.org/10.1016/bs.aecr.2018.01.001

Lukács G (1971) History and Class Consciousness (Translated by Rodney Livingstone). Talgarth: The Merlin Press Ltd.

Mann DG, Vanormelingen P (2013) An Inordinate Fondness? The Number, Distributions, and Origins of Diatom Species. Journal of Eukaryotic Microbiology 60: 414–420. https://doi.org/10.1111/jeu.12047

Mann DG, Sato S, Trobajo R, Vanormelingen P, Souffreau C (2010) DNA barcoding for species identification and discovery in diatoms. Cryptogamie, Algologie 31: 557–577.

Poikane S, Birk S, Böhmer J, Carvalho L, de Hoyos C, Gassner H, Hellsten S, Kelly M, Solheim AL, Olin M, Pall K, Phillips G, Portielje R, Ritterbusch D, Sandin L, Schartau A-K, Solimini AG, van den Berg M, Wolfram G, van de Bund W (2015) A hitchhiker's guide to European lake ecological assessment and intercalibration. Ecological Indicators 52: 533–544. https://doi.org/10.1016/j.ecolind.2015.01.005

Rimet F, Bouchez A (2012) Life-forms, cell-sizes and ecological guilds of diatoms in European rivers. Knowledge and Management of Aquatic Ecosystems 406(01). https://doi.org/10.1051/kmae/2012018

Rimet F, Franc A, Gusev E, Kahlert M, Kelly M, Kulikovskiy M, Laval V, Maltsev Y, Mann D, Pfannkuchen M, Trobajo R, Vasselon V, Zimmermann J, Bouchez A (2019) Diat.barcode, an open-access barcode library for diatoms: new insights. Scientific Reports 9, Article number: 15116. https://doi.org/10.1038/s41598-019-51500-6

Ruck EC, Nakov T, Alverson A, Theriot EC (2016) Phylogeny, ecology, morphological evolution, and reclassification of the diatom orders Surirellales and Rhopalodiales. Molecular Phylogenetics and Evolution 103: 155–171. https://doi.org/10.1016/j.ympev.2016.07.023

Tapolczai K, Bouchez A, Stenger-Kovács C, Padisák J, Rimet F (2016) Trait-based ecological classifications for benthic algae: review and perspectives. Hydrobiologia 776: 1–17. https://doi.org/10.1007/s10750-016-2736-4

Vasselon V, Bouchez A, Rimet, F, Jacquet S, Trobajo R, Corniquel M, Tapolczai K, Domaizon I (2017) Avoiding quantification bias in metabarcoding: application of a cell 1 biovolume correction factor in diatom molecular biomonitoring. Methods in Ecology and Evolution 9: 1060–1069. https://doi.org/10.1111/2041-210X.12960

Weigand H, Beermann AJ, Čiampor F, Costa FO, Csabai Z, Duarte S, Geiger MS, Grabowski M,Rimet F, Rulik B, Strand M, Szucsich N, Weigand AM, Willassen E, Wyler SA, Bouchez A, Borja A, Čiamporová-Zaťovičová Z, Ferreira S, Dijkstra KD, Eisendle U, Freyhof J, Gadawski P, Graf W, Haegerbaeumer A, van der Hoorn BB, Japoshvili B, Keresztes L, Keskin E, Leese F, Macher J, Mamos T, Paz G, Pešić V, Pfannkuchen DM, Pfannkuchen MA, Price BW, Rinkevich B, Teixeira MAL, Várbíró G, Ekrem T (2019) DNA barcode reference libraries for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. Science of the Total Environment 678: 499–524. https://doi.org/10.1016/j.scitotenv.2019.04.247

Willby N (2011) From metrics to Monet: the case for an ecologically-meaningful guiding image. Aquatic Conservation: Marine and Freshwater Ecosystems 21: 601–603. https://doi.org/10.1002/aqc.1233

Woodward G, Gray C, Baird DJ (2013) Biomonitoring for the 21st Century: New perspectives in an age of globalisation and emerging environmental threats. Limnetica 32: 159–174.